Proceedings

10th IEEE International Conference on Data Mining

Sydney, Australia
14–17 December 2010

Edited by
Geoffrey I. Webb, Bing Liu, Chengqi Zhang, Dimitrios Gunopulos, and Xindong Wu

CPS Conference Publishing Services

Los Alamitos, California

Washington • Tokyo

IEEE computer society

IEEE Computer Society Order Number P4256
BMS Part Number CFP10278-CDR
ISBN 978-0-7695-4256-0

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
http://computer.org/cspress
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: <reprints@computer.org>

Editorial production by Lisa O'Conner

IEEE Computer Society
***Conference Publishing Services* (CPS)**
http://www.computer.org/cps

2010 IEEE International Conference on Data Mining

ICDM 2010

Table of Contents

Keynote Abstracts

Regular Papers

Short Papers

Tutorials

Message from the Program Committee Co-Chairs

We welcome you to the tenth annual IEEE International Conference on Data Mining (ICDM-2010), held this year in Sydney, Australia. This year saw further progression of the trajectory that has seen ICDM become a truly international, highly selective and leading conference in the field of data mining, where only the best research results are chosen to be presented at the conference. Over the years, the ICDM conference series has not only served as a leading forum for researchers to present their work and to exchange ideas with one another, but also as a meeting place for practitioners to discuss their applications with researchers and to gain new knowledge from the latest research results. ICDM continues to play a central role in data mining's growth from laboratory research to wide spread applications across numerous fields of endeavour.

Building on a steady upward trajectory, this year the conference received a record of 797 submissions from 53 countries, demonstrating that our field continues to grow and flourish. The 33 vice-chairs were drawn from 12 countries. The 317 program committee members came from 30 countries. The program committee was assisted by a further 96 reviewers drawn from 14 countries.

The submitted papers spanned a broad spectrum of current topics in data mining. Only 72 papers were accepted for regular presentation at the conference, an acceptance rate of only 9%. A further 83 papers were accepted for short presentations, a total acceptance rate of 19%.

It is a massive process to undertake almost 2,700 reviews, discuss and reconcile differences in opinions between those reviews, and winnow the final set of selected papers. The community owes a debt of gratitude to the vice chairs, program committee members and external reviewers. Without their dedicated voluntary work, this conference would not be possible. We are also deeply appreciative of Juzhen Dong who maintains the Cyberchair paper submission and reviewing process. We are indebted to her for her prompt and effective implementation of several new features at short notice.

In addition to the technical papers, the program includes invited talks by Christos Faloutsos, Geoff McLachlan and Xindong Wu, a panel on the Top-10 Data Mining Case Studies, the fourth data mining contest, 4 tutorials, 18 workshops, exhibits and demos, all capped off by a stimulating social program that will provide for the informal interactions that are the life blood of a scientific discipline.

We thank the honorary chair Ramamohanarao Kotagiri, the conference co-chairs Chengqi Zhang and Dimitrios Gunopulos, the steering committee chair Xindong Wu and the steering committee for their critical role in overseeing the conference and ensuring its steady progression from success to success.

The workshop co-chairs Wynne Hsu and Wei Fan, the tutorial co-chairs Sanjay Chawla and Myra Spiliopoulou, the panels chair Hillol Kargupta, the exhibits and demos co-chairs Bart Goethals and Yucel Saygin, and the contest chair Demetris Zeinalipour, awards committee chair Wei Wang, student travel awards committee chair Wei Ding, and all their many committee members have all played vital roles in the development of a rounded, exciting and stimulating program. They are deeply deserving of the community's great appreciation for their dedicated voluntary service.

At the direction of the Steering Committee we undertook an experiment to assess the inter-rater reliability of the ICDM review process. To this end we duplicated ten papers and had them reviewed in parallel. Unfortunately this could not be integrated into the automated review assignment process, so reviewers were added after all other reviewer assignments had been completed. The outcome was that the duplicates had a far higher acceptance rate than the originals. We believe this is the result of poor reviewer assignment and the lesson learned is that good reviewer assignment is critical. Due to the poor quality of the reviewing for the duplicated papers, the duplicate reviews were not taken

into account in the final acceptance and rejection process. More detailed analysis of the outcomes will be undertaken, and this will be presented at the Community Meeting.

We hope that you find the breadth and depth of this year's outstanding technical program as stimulating and inspiring as we have.

Geoff Webb and Bing Liu,

ICDM-2010 Program Committee Co-Chairs